

Référence d'offre de stage : S22024 ([English description](#) : page 2 ; [Description Française](#) : page 1)

Stage – Master 2 en apprentissage automatique et vision par ordinateur

« Détection d'objets en few-shot basée sur un cross-transformer »

Ce stage s'inscrit dans le cadre du LabCom IRISER, Laboratoire Commun en « Intelligence, Reconnaissance, Surveillance Réactive » (<https://www-l2ti.univ-paris13.fr/iriser/>) financé par l'ANR. Le LabCom IRISER vise à proposer, à maîtriser de bout en bout le comportement et les performances des systèmes intelligents voire embarqués destinés à la vision artificielle pour l'analyse rapide et automatisée des images/vidéos (de très grandes tailles, multispectrales géoréférencées de hautes résolutions) capturées à bord des avions de COSE en s'appuyant sur des stratégies de traitement de l'information visuelle et d'apprentissage automatique. Nous souhaitons recruter un stagiaire de master en recherche et développement en vision par ordinateur et apprentissage machine sur une durée de 5 mois.

Description succincte du projet scientifique

La détection d'objets dans un contexte few-shot (FSOD) a pour objectif de détecter des objets dans une image de requête en utilisant seulement quelques exemples d'entraînement. Cette approche s'inspire du système visuel humain capable d'assimiler rapidement de nouveaux concepts à partir d'un nombre limité d'exemples. Le défi central consiste à apprendre rapidement des modèles de détection d'objets, même avec un faible volume de données d'entraînement. Un tel modèle serait alors apte à détecter des objets dans des images nouvelles. Cette tâche s'avère particulièrement complexe, surtout pour les approches reposant sur l'apprentissage profond. Ces approches exigent généralement des milliers d'exemples d'entraînement. De plus, elles sont susceptibles de présenter un surajustement dans des contextes de données limitées. Les méthodes développées suivent principalement un paradigme d'apprentissage en deux étapes, visant à transférer les connaissances acquises à partir des classes de base pour assister la détection d'objets dans le contexte des nouvelles classes (en few-shot). Les architectures, pouvant être globalement classées en deux catégories : les méthodes basées sur une seule branche et les méthodes basées sur deux branches.

Dans ce stage, il s'agira d'étudier le nouveau modèle basé sur le Fully Cross-Transformer (FCT) [1]. Ce dernier intègre un cross-transformer à la fois dans le réseau de caractéristiques et dans la tête de détection. Une évaluation des performances du modèle de détection d'objets sur des images aériennes (DOTA, DIOR) est demandée. Une importance particulière sera accordée aux petits objets.

Mots clefs : Object Detection, Zero-Shot Learning, Transformers, Deep learning, Aerial Images.

Référence :

[1] Guangxing Han, Jiawei Ma, Shiyuan Huang, Long Chen, Shih-Fu Chang, Few-Shot Object Detection with Fully Cross-Transformer, CVPR 2022.

Profil du candidat(e) : Vous êtes issu(e) d'une formation (Master 2) en sciences de données (traitement du signal des images, informatique avec une dominante en intelligence artificielle/apprentissage/machine learning, mathématiques appliquées) et vous êtes à l'aise en développement informatique (Python/ framework Pytorch, Tensorflow). Vous maîtrisez le traitement d'images et les techniques d'intelligence artificielle (Deep learning ...). Vous êtes à l'aise en Anglais aussi bien à l'écrit qu'à l'oral. Vous êtes motivé (e) et curieux (se) de nature, dans ce cas venez participer à la création du LabCom IRISER.

Durée du stage : 5 mois

Date de début prévue : 15/03/2024 -- **Date limite de dépôt** : 29/02/2024

Contact : Fangchen FENG, Anissa MOKRAOUI,

Courriel : fangchen.feng@univ-paris13.fr ; anissa.mokraoui@univ-paris13.fr

Dossier de candidature : CV, lettre de motivation, relevés de notes, au moins une lettre de recommandation (merci d'indiquer la référence d'offre de stage dans le sujet de votre courriel).

Référence d'offre de stage : S22024 (English version)

Internship – Master's Degree in Machine Learning and Computer Vision

« Few-Shot Object Detection using a Cross-Transformer »

This internship is part of the LabCom IRISER, a Common Laboratory in "Intelligence, Reconnaissance, Reactive Surveillance" (<https://www-l2ti.univ-paris13.fr/iriser/>) funded by the ANR. The LabCom IRISER aims to propose end-to-end control over the behavior and performance of intelligent or embedded systems intended for artificial vision for the rapid and automated analysis of images/videos (of very large sizes, georeferenced multispectral images of high resolution) captured aboard COSE aircraft, relying on strategies for visual information processing and machine learning. We are looking to recruit a master's research and development intern in computer vision and machine learning for a duration of 5 months.

Brief Description of the Scientific Project:

Object detection in a few-shot context (FSOD) aims to detect objects in a query image using only a few training examples. This approach draws inspiration from the human visual system's ability to quickly assimilate new concepts from a limited number of examples. The central challenge is to rapidly learn object detection models, even with a small volume of training data, making the model capable of detecting objects in new images. This task proves particularly complex, especially for deep learning-based approaches, which typically require thousands of training examples. Additionally, they are prone to overfitting in limited data contexts.

The developed methods mainly follow a two-step learning paradigm, aiming to transfer knowledge acquired from base classes to assist in object detection in the context of new classes (few-shot). Architectures can be broadly classified into two categories: methods based on a single branch and methods based on two branches.

In this internship, the focus will be on studying the novel model based on the Fully Cross-Transformer (FCT) [1]. The FCT integrates a cross-transformer both in the feature network and in the detection head. An evaluation of the object detection model's performance on aerial images (DOTA, DIOR) is required, with particular attention given to small objects.

Key Words : Object Detection, Zero-Shot Learning, Transformers, Deep learning, Areal Images.

Reference :

[1] Guangxing Han, Jiawei Ma, Shiyuan Huang, Long Chen, Shih-Fu Chang, Few-Shot Object Detection with Fully Cross-Transformer, CVPR 2022.

Candidate Profile:

You hold a Master's degree (Master 2) in data science (image signal processing, computer science with a focus on artificial intelligence/machine learning, applied mathematics), and you have proficiency in computer programming (Python/PyTorch framework, Tensorflow). You possess expertise in image processing and artificial intelligence techniques (Deep learning, etc.). You are comfortable with both written and spoken English. If you are motivated and naturally curious, we invite you to join the creation of LabCom IRISER.

Internship Duration: 5 months Scheduled

Start Date: 15/03/2024 -- **Application Deadline:** 29/02/2024

Contact: Fangchen FENG, Anissa MOKRAOUI

Email: fangchen.feng@univ-paris13.fr ; anissa.mokraoui@univ-paris13.fr

Application Documents: CV, cover letter, transcripts, at least one letter of recommendation (please indicate the internship offer reference in the subject of your email).